

PAPER • OPEN ACCESS

Patch-based anisotropic diffusion scheme for fluorescence diffuse optical tomography—part 1: technical principles

To cite this article: Teresa Correia and Simon Arridge 2016 *Phys. Med. Biol.* **61** 1439

View the [article online](#) for updates and enhancements.

Related content

- [Patch-based anisotropic diffusion scheme for fluorescence diffuse optical tomography—part 2: image reconstruction](#)
Teresa Correia, Maximilian Koch, Angélique Ale et al.
- [Maximum likelihood estimation-based denoising of magnetic resonance images](#)
Jeny Rajan, Ben Jeurissen, Marleen Verhoye et al.
- [Denoising human cardiac diffusion tensor magnetic resonance images](#)
L J Bao, Y M Zhu, W Y Liu et al.

Patch-based anisotropic diffusion scheme for fluorescence diffuse optical tomography—part 1: technical principles

Teresa Correia and Simon Arridge

Centre for Medical Imaging Computing, Department of Computer Science, University College London, Gower Street, London WC1 E6BT, UK

E-mail: tcorreia@cs.ucl.ac.uk

Received 6 May 2014, revised 29 September 2015

Accepted for publication 30 November 2015

Published 25 January 2016



CrossMark

Abstract

Fluorescence diffuse optical tomography (fDOT) provides 3D images of fluorescence distributions in biological tissue, which represent molecular and cellular processes. The image reconstruction problem is highly ill-posed and requires regularisation techniques to stabilise and find meaningful solutions. Quadratic regularisation tends to either oversmooth or generate very noisy reconstructions, depending on the regularisation strength. Edge preserving methods, such as anisotropic diffusion regularisation (AD), can preserve important features in the fluorescence image and smooth out noise. However, AD has limited ability to distinguish an edge from noise. In this two-part paper, we propose a patch-based anisotropic diffusion regularisation (PAD), where regularisation strength is determined by a weighted average according to the similarity between patches around voxels within a search window, instead of a simple local neighbourhood strategy. However, this method has higher computational complexity and, hence, we wavelet compress the patches (PAD-WT) to speed it up, while simultaneously taking advantage of the denoising properties of wavelet thresholding. The proposed method combines the nonlocal means (NLM), AD and wavelet shrinkage methods, which are image processing methods. Therefore, in this first paper, we used a denoising test problem to analyse the performance of the new method. Our results show that the proposed PAD-WT method provides better results than the AD or NLM methods alone. The efficacy of the method for fDOT image reconstruction problem is evaluated in part 2.



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Keywords: fluorescence diffuse optical tomography, image reconstruction, anisotropic diffusion, nonlocal means, structural information, multimodal imaging, regularisation

(Some figures may appear in colour only in the online journal)

1. Introduction

Fluorescence diffuse optical tomography (fDOT) is an optical imaging modality that provides three-dimensional (3D) images of fluorescent source distributions inside biological tissue (Ntziachristos 2006). The image reconstruction in fDOT is very challenging due to ill-posedness of the inverse problem, which is a consequence of the multiple scattering nature of biological tissues. To overcome this problem, regularisation methods are commonly used to stabilise the solution. Usually, a quadratic (L2 norm) regularisation or penalty term is added to the least squares objective function to enforce smoothness in the reconstructed image (Hansen 1998, Zacharopoulos *et al* 2009). In Correia *et al* (2011) we proposed to use a nonlinear anisotropic diffusion regularisation method, which has the ability to smooth out noise while preserving edges in images, and showed that spatial localisation and size of fluorescence inclusions can be accurately estimated. In this method, a reconstruction step alternates with the regularisation step, i.e. a nonlinear anisotropic diffusion (AD) filtering step (Perona and Malik 1990). The AD method is a well-known and widely used image processing technique that provides satisfactory noise suppression and edge enhancement results.

Buades *et al* (2005a) proposed the nonlocal means method (NLM), an image processing algorithm that surpasses the AD method. In the NLM method, pixels are averaged according to similarity between patches around pixels within a search window W , instead of a simple averaging strategy within a local neighbourhood. NLM exploits the redundancy of information within an image, by assuming that patches from different regions contain similar patterns and averaging them effectively reduces noise.

The NLM has superior denoising performance than local-based methods, but at the expense of higher computational complexity. Several methods have been proposed to accelerate the NLM without loss of denoising performance, such as a block-wise implementation of the NLM (Buades *et al* 2005b, Coupé *et al* 2008), preselection of similar patches based on mean and average gradient of patches (Mahmoudi and Sapiro 2005) or mean and variance (Coupé *et al* 2008), arrangement of patches in a cluster tree (Brox *et al* 2008), singular value decomposition (SVD) (Orchard *et al* 2008) and principal component analysis (PCA) (Tasdizen 2008). NLM using PCA compressed patches not only reduces the computational time, but also improves the denoising performance (Tasdizen 2008).

Another popular denoising/compression method is the wavelet shrinkage, which is based on the process of transforming an image into the wavelet domain (Donoho and Malik 1994, Donoho 1995, Donoho and Johnstone 1995). Important information is encoded by large wavelet coefficients, whereas most of small coefficients represent noise and can be removed using thresholding techniques. The denoised image is obtained by transforming the thresholded coefficients into the image domain.

The NLM method has been successfully used in medical imaging, for example, as a denoising technique (Coupé *et al* 2012, Chen *et al* 2012, Dutta *et al* 2013, Chan *et al* 2014) and in image reconstruction as a regularisation method with (Chen *et al* 2008, Chun *et al* 2012, Wang and Qi 2012, Nguyen and Lee 2013). Here, motivated by the effectiveness of the NLM method in medical imaging applications, we propose a modification to our nonlinear anisotropic

diffusion regularisation method for fDOT image reconstruction (Correia *et al* 2011). In this work, instead of the previous local strategy, we consider a patch-based approach where we use a robust edge-preserving potential function. We refer to this method as patch-based anisotropic diffusion (PAD). Moreover, we propose to reduce patch dimensions using wavelet compression, not only to reduce computational complexity but also to increase the robustness of the method to noise. Therefore, the new method combines the advantages of NLM, AD and wavelet shrinkage methods. We refer to this method as patch-based anisotropic diffusion with wavelet patch compression (PAT-WT). Since our new method is based on the NLM method, which was initially proposed for image denoising, we begin by assessing the performance of the PAD(-WT) method using a 2D denoising test problem (an ill-posed inverse problem). In the second part of this two-part paper, we study the performance of the PAD-WT method in fDOT image reconstruction as a regulariser.

2. Methods

Consider a noisy image of the form:

$$f_{\text{noise}} = f_{\text{true}} + \varepsilon, \quad (1)$$

where f_{true} is the true image and ε is the noise. Denoising algorithms aim at removing noise from the observed noisy image f_{noise} , returning an image as close as possible to f_{true} . The NLM and AD are examples of widely used image denoising methods.

2.1. Anisotropic diffusion

In the image space $f \in \mathbb{R}^2$ the discretised anisotropic diffusion method is given by the following iterative scheme (Perona and Malik 1990):

$$f^{[k+1]} = f^{[k]} + \frac{\tau}{|\mathbf{W}(i)|} \sum_i \sum_{j \in \mathbf{W}(i)} g(|f_j^{[k]} - f_i^{[k]}|) (f_j^{[k]} - f_i^{[k]}) \quad (2)$$

where k is the iteration number, $\tau \in [0, 1]$ is the time step, $\mathbf{W}(i)$ is the set of neighbours of pixel i with constant dimension, $|\mathbf{W}(i)|$ is the number of neighbours (here $|\mathbf{W}(i)| = 4$, 2 neighbours per direction) and g is an edge-preserving function such as Huber, Tukey, Total Variation, Perona–Malik, etc (Correia *et al* 2011). Here, we use the Perona–Malik function defined by $g(|f_j - f_i|) = 1 / \left(1 + \left(\frac{|f_j - f_i|}{T} \right)^2 \right)$, for a pixel $j \in \mathbf{W}(i)$. The parameter T is the threshold and can be selected using the normalised cumulative histogram (NCH) of the gradient (Correia *et al* 2011). This method is commonly used in edge detection problems. The NCH indicates the probability φ of a gradient taking on a value less than or equal to the value X that the bin represents, i.e. $\varphi(|f_j - f_i| \leq X)$. It increases monotonically and the smoothness/sharpness of the curve indicates how smooth/sharp the edges are. The threshold can be calculated from the NCH by setting the threshold at, for example, 90 percent.

2.2. Patch-based anisotropic diffusion

We propose the following modification to the previous method:

$$f^{[k+1]} = f^{[k]} + \tau \sum_i \sum_{j \in \mathbf{W}(i)} w_{i,j}(f^{[k]}) g(|f_j^{[k]} - f_i^{[k]}|) (f_j^{[k]} - f_i^{[k]}), \quad (3)$$

where w_{ij} is the weight that measures the similarity between two neighbourhoods (patches with fixed size $P = N \times N$) centred at pixel i and j , in a search window $W = (2W + 1) \times (2W + 1)$ centred at pixel i , where W is the maximum distance from pixel i (see figure 1 for an example). The weight w_{ij} is defined as:

$$w_{ij}(f) = \frac{1}{C(i)} \exp\left\{-\frac{\mathcal{D}(f(\mathcal{N}_i) - f(\mathcal{N}_j))}{h^2}\right\}, \quad (4)$$

where \mathcal{N}_i and \mathcal{N}_j are the neighbourhoods centred at pixel i and j , respectively. The parameter h controls the exponential decay and $C(i)$ is a normalising constant given by:

$$C(i) = \sum_j \exp\left\{-\frac{\mathcal{D}(f(\mathcal{N}_i) - f(\mathcal{N}_j))}{h^2}\right\} \quad (5)$$

and

$$\mathcal{D}(f(\mathcal{N}_i) - f(\mathcal{N}_j)) = \sum_{p=1}^{N \times N} (f_{ip} - f_{jp})^2, \quad (6)$$

where f_{ip} and f_{jp} represent the p element of the patches $f(\mathcal{N}_i)$ and $f(\mathcal{N}_j)$, respectively. Therefore, the weights w_{ij} are large when patches are similar and small when they are much different.

2.3. Nonlocal means

The new patch-based anisotropic diffusion method is based on the nonlocal means denoising method (Buades *et al* 2005a, 2005b):

$$f_{\text{NLM}} = \sum_i \sum_{j \in W(i)} w_{i,j}(f) f. \quad (7)$$

NLM can be considered as a one step of a fixed point iteration (Brox *et al* 2008), and denoising results can be improved by performing a few iterations:

$$f^{[k+1]} = \sum_i \sum_{j \in W(i)} w_{i,j}(f^{[k]}) f^{[k]}. \quad (8)$$

The iterative NLM method computes a weighted average of pixel intensities at each iteration. In our proposed method, pixel intensities are replaced by $g(|f_j^{[k]} - f_i^{[k]}|)(f_j^{[k]} - f_i^{[k]})$. Note that the patch-based anisotropic diffusion (PAD) method becomes the local AD when $N = 1$ and $W = 1$.

2.4. Patch compression

NLM methods have superior denoising performance than local methods, but at the expense of computational complexity. Principal component analysis (PCA) has been used to reduce data dimensionality and speed up the NLM method (Tasdizen 2008). Furthermore, accuracy of the solution was shown to improve since principal components that represent noise, i.e. with small variance, are removed. We propose to use a wavelet transform (WT) compression method instead. In this method, the high frequency coefficients are removed, preserving the main information of the data, while removing noise. Compression is applied to the patches, which are previously converted into single dimension arrays to simplify and speed up computations.

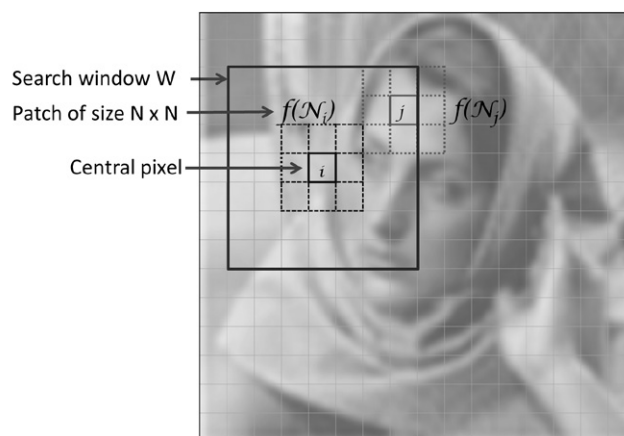


Figure 1. Similar to the NLM method, the patch-based anisotropic diffusion method is based on patch similarity. A square search window W is defined surrounding pixel i . The method estimates pixel i by taking a weighted average of all pixels j within W . The weight w_{ij} is calculated based on the similarity between two neighbourhoods or patches $f(\mathcal{N}_i)$ and $f(\mathcal{N}_j)$, centred around pixel i and j , respectively. Similar patches give larger weights. In this 2D example, patches have dimensions $P = 3 \times 3$, hence, patch length is $N = 3$, and the window $W = 7 \times 7$ or $\mathcal{W} = 3$, meaning that the maximum pixel distance from pixel i is 3 pixels.

The patch-based anisotropic diffusion method with wavelet patch compression (PAD-WT) is our proposed method.

2.4.1. Principal component analysis. PCA is a statistical method widely used in data analysis and compression. PCA identifies similarities and differences in the data, by representing the signal in terms of a set of new orthogonal variables called principal components. The first and last components have the largest and smallest variance, respectively.

Consider a patch $f(\mathcal{N}_i)$ with P elements, the mean of its elements \bar{f}_i and covariance $\text{cov}(f(\mathcal{N}_i)) = (\sum_{p=1}^P (f_{ip} - \bar{f}_i)(f_{ip} - \bar{f}_i)^T) / (P - 1)$. The eigendecomposition of the covariance matrix is $\text{cov}(f(\mathcal{N}_i)) = U \Lambda U^{-1} = U \Lambda U^T$, where Λ is a diagonal matrix whose elements are the eigenvalues and the columns of the orthonormal matrix U are the eigenvectors of the covariance matrix. The principal components are the eigenvectors ordered by descending magnitude of the corresponding eigenvalues. The first component is the principal component and represents the largest variance of the data. The dimensionality of the data can be reduced to d by projecting the patch $f(\mathcal{N}_i)$ onto the new orthogonal basis, $p_i = U_d^T f(\mathcal{N}_i)$, where the d largest eigenvectors are kept, i.e. the first d columns of U , and the less significant components are discarded.

2.4.2. Wavelet transform. The fast wavelet transform (FWT) decomposes a signal or function into different frequency subbands (Mallat 2008). It uses a low-pass filter H and a high-pass filter G to obtain the approximation ϖ_c and detail ϖ_d coefficients. The approximation coefficients represent the approximation of the signal at a resolution 2^r , where r is an integer that specifies the resolution level. The detail coefficients contain the details of the original signal, i.e. the high-frequency information. For a signal $f(n) = \varpi_i(r + 1, n)$, with n samples and at a starting scale $r + 1$:

$$\varpi_c(r, n) = H(-n) * \varpi_c(r + 1, n) \downarrow 2, \quad (9)$$

$$\varpi_d(r, n) = G(-n) * \varpi_d(r + 1, n) \downarrow 2. \quad (10)$$

Thus, the FWT consists of a convolution, followed by downsampling by a factor of 2 ($\downarrow 2$), i.e. keeping the even index samples. The wavelet transform can be implemented as a decomposition filter bank, where the initial signal goes through a series of filters. Note that the filters are related to each other by $G(n) = (-1)^n H(1 - n)$.

Data compression is used to reduce the dimensions of the patches for computational efficiency. If a two-level 1D FWT is applied to patch $f(\mathcal{N}_i)$, the obtained wavelet coefficients are $\varpi_i = \{\varpi_c(r - 1, n), \varpi_d(r - 1, n), \varpi_d(r, n)\}$. The vectorised patch is compressed by keeping the ℓ largest coefficients, which contain most of the relevant information. Thus, the weights in equations (3) and (4) are calculated using:

$$\mathcal{D}(f(\mathcal{N}_i) - f(\mathcal{N}_j)) = \sum_{p=1}^{\ell} (\varpi_{ip} - \varpi_{jp})^2, \quad (11)$$

where ϖ_{ip} and ϖ_{jp} represent the p wavelet coefficient of the wavelet transformed patches $f(\mathcal{N}_i)$ and $f(\mathcal{N}_j)$, respectively.

3. Evaluation

We used an image denoising problem to test the hypotheses that the proposed method is superior to NLM, anisotropic diffusion filtering and that WT is an efficient patch compression method. We used several images corrupted with Gaussian noise in the denoising test. First, we analysed the performance of the PAD-WT method for different wavelet types. Finally, we solved the denoising problem using the following iterative methods: (1) NLM (see equation (8)), (2) NLM with patch compression using PCA (NLM-PCA), (3) NLM with patch compression using wavelets (NLM-WT), (4) AD or local PAD (see equation (2)), (5) PAD (see equation (3)), (6) PAD with patch compression using PCA (PAD-PCA) and (7) PAD-WT.

We considered two tests: Test (1) using a test image (figure 2) corrupted by Gaussian noise with different noise levels, ranging from 10% to 50% in steps of 10%, and Test (2) several standard test images (figure 2) with the same noise level (20% Gaussian noise). All the images had 320×320 pixels.

We assessed the denoising quality with a widely used figure of merit in image processing, the peak signal to noise ratio (PSNR) between the original image f_{true} and the denoised image f : $\text{PSNR} = 10 \log_{10}(\max(f_{\text{true}})/\text{MSE})$, where $\text{MSE} = \frac{1}{\text{NP}} \sum_{i=1}^{\text{NP}} (f_{\text{true},i} - f_i)^2$ and NP is the total number of pixels in the image.

We solved the denoising problem for 20 noise realisations for each noise level and image. For all the methods, the iterative process was automatically stopped when $\text{PSNR}^{k+1} < \text{PSNR}^k$. We averaged the resulting PSNRs and calculated the corresponding standard deviations.

As mentioned previously, the parameters T and h can be calculated from the NCH by setting the threshold at a certain percentage. The PSNR was calculated for different possible combinations of percentages (for NLM-based methods we only calculated h). The parameters that resulted in the highest PSNR were the ones used to generate the results presented here. The time step was set to $\tau = 1$ for the PAD and AD methods. All methods were run on a linux PC with an Intel Xeon E5-2665 CPU @ 2.4 GHz using Matlab Mex-files.



Figure 2. Images used in Test 1 (left) and Test 2 (from left to right: Barbara, cameraman, CT and MRI).

4. Results and discussion

4.1. Wavelet types

We used Test 1 (left image in figure 2 with different noise levels) to identify a suitable wavelet type for the patch compression. The wavelet types tested were (Salomon and Motta 2009): Haar, Vaidyanathan, Daubechies of order 2, 4 and 6, Battle-Lemarié of order 1, 3 and 5, and Coiflet of order 1, 3 and 5. The patch and search window sizes of the PAD-WT method were set to $N = 3$ and $\mathcal{W} = 5$, respectively. The largest $\ell = \lceil (N \times N)/4 \rceil$ wavelet coefficients are kept, where $\lceil \cdot \rceil$ denotes the ceil function. For a $N = 3$ patch, $\ell = 3$, which is considered to be the maximum compression for this patch size.

Table 1 shows the averaged PSNRs (and standard deviations) obtained using the proposed method with different wavelet types to remove noise from an image corrupted with 5 different noise levels. The best results were obtained using Daubechies wavelets of order 4. Therefore, Daubechies wavelets of order 4 were used to compress the patches in the following analysis.

4.2. Comparison between denoising methods

First, we used Test 1 to compare the iterative NLM method with PAD and their local versions. As mentioned previously, the methods are considered local when $N = 1$ and $\mathcal{W} = 1$. We compared the performance of the previous methods to their equivalent with patch compression, using PCA or WT.

We calculated the averaged PSNRs and standard deviations of images obtained using the NLM and PAD methods, with and without patch compression, for window sizes $\mathcal{W} = \{1, 3, 5, 7, 15\}$ and patches of length $N = \{3, 7\}$. We retained a total of $\ell = \lceil (N \times N)/4 \rceil$ wavelet coefficients for the NLM(PAD)-WT method. For the NLM(PAD)-PCA method, we kept 1/2 of the principal components.

Table 2 shows the denoising results for Test 1 obtained using equation (8) with $N = 1$ and $\mathcal{W} = 1$ (local filtering), NLM, NLM-PCA and NLM-WT methods. Table 3 shows the results obtained using AD, PAD-PCA, PAD-WT, and the respective elapsed time for one iteration. Results show that NLM-based methods perform better than local filtering. This test also shows that the PAD method performs better than the conventional NLM or AD. Similarly, the PAD-WT shows superior performance compared to the NLM-WT method. Better results were obtained using PAD(NLM)-WT compared to PAD(NLM)-PCA. The PAD-WT method provides results close to those of PAD without patch compression for lower levels of noise and gives better results for higher noise levels ($>30\%$), with the advantage of being faster.

These results show that the size of the search window affects the performance of the PAD(NLM) method. This test suggest that for the PAD-WT method the optimal window and patch sizes are $\mathcal{W} = 7$ and $N = 3$, respectively. The results obtained using $\mathcal{W} = 5$ are almost

Table 1. Denoising PSNR results for a test image with various noise levels obtained using PAD-WT with different wavelet types.

	10%	20%	30%	40%	50%
Haar	36.01 ± 0.11	31.39 ± 0.12	28.95 ± 0.11	26.93 ± 0.14	25.19 ± 0.16
Vaidyanathan	35.99 ± 0.10	31.59 ± 0.12	28.81 ± 0.13	26.60 ± 0.16	24.82 ± 0.13
Daubechies 2	36.07 ± 0.09	31.61 ± 0.12	28.78 ± 0.13	26.96 ± 0.18	25.21 ± 0.16
Daubechies 4	36.22 ± 0.15	31.82 ± 0.15	29.22 ± 0.14	27.34 ± 0.16	25.53 ± 0.19
Daubechies 6	36.17 ± 0.12	31.69 ± 0.11	29.07 ± 0.17	26.93 ± 0.18	25.18 ± 0.23
Battle-Lemarié 1	35.51 ± 0.08	31.47 ± 0.07	28.99 ± 0.18	26.95 ± 0.22	25.33 ± 0.19
Battle-Lemarié 3	34.95 ± 0.08	30.25 ± 0.09	27.52 ± 0.12	25.25 ± 0.19	23.69 ± 0.11
Battle-Lemarié 5	36.16 ± 0.09	31.67 ± 0.14	29.02 ± 0.14	26.93 ± 0.21	25.12 ± 0.17
Coiflet 1	34.85 ± 0.13	30.40 ± 0.10	27.76 ± 0.10	25.64 ± 0.12	24.14 ± 0.10
Coiflet 3	30.58 ± 0.07	26.20 ± 0.08	23.89 ± 0.08	22.36 ± 0.09	21.25 ± 0.06
Coiflet5	31.60 ± 0.06	26.85 ± 0.07	23.89 ± 0.08	22.43 ± 0.09	21.07 ± 0.10

Note: The largest PSNR values are marked.

as good, with the advantage of being approximately two times faster to compute. For a window of size $\mathcal{W} = 10$ and patch $N = 7$, as used by Buades *et al* (2005b), the method is slower and the PSNR values are not higher.

In this study, we did not intend to perform an extensive study on the optimal number of wavelet coefficients that results in the highest PSNR value. Nevertheless, we tested other ℓ values and there were no notable changes in the results (not shown).

Tasdizen (2008) found that, for a patch of size $N = 7$, choosing $d = 6$ yields the highest PSNR. The performance of the NLM(PAD)-PCA for $d = 6$ and $d = \lceil (N \times N) t \rceil$, where $t = \{3/8, 1/4\}$, was also evaluated. The PSNRs values (not shown) do not vary significantly, and hence, the conclusions regarding the performance of the denoising methods remain the same. Nevertheless, for consistency, d was kept fixed since ℓ was also kept constant.

Figure 3 shows the test image (Test 1) contaminated with different levels of noise and the best denoising results obtained using the PAD, PAD-PCA and PAD-WT methods. The proposed PAD-WT method is qualitatively very similar to the PAD and performs much better than the PAD-PCA method. However, note that the image background can be recovered using PAD-PCA if different T and h parameters are used, but at the expense of decreased PSNR values. Both PAD and PAD-WT demonstrate high denoising performance for noise levels up to 20%. For higher noise levels, low contrast features cannot be recovered. However, for noise levels up to 40% the quality of the denoised images is still quite high. Iteration times are faster when wavelet patch compression is used compared to other PAD-based methods (table 3).

Test 2 was used to further evaluate the performance of the PAD-WT method. Denoising results for standard test images using PAD, PAD-PCA and PAD-WT are shown in table 4 and figure 4. For each method we used the window and patch sizes that gave the best Test 1 results for a noise level of 20%. The PAD-WT method returns slightly lower PSNRs than the PAD method for the Barbara and cameraman images, but higher values for the CT and MRI images. This denoising study suggests that the proposed method is suitable for medical imaging applications.

One of the advantages of using the AD method is the computational speed. However, NLM-based methods can easily be parallelised and implemented to exploit GPU (graphics processing units) acceleration (Cuomo *et al* 2014) to greatly reduce the execution times. Alternatively, to reduce the computational times, a semi-implicit scheme can be used, instead of the explicit discretisation used here, so that τ can be relatively large without causing

Table 2. Denoising PSNR results for a test image with various noise levels (Test 1) obtained using local filtering, NLM, NLM-PCA and NLM-WT.

	\mathcal{W}	10%	20%	30%	40%	50%	P
Local	1	29.18 \pm 0.05	26.13 \pm 0.06	23.87 \pm 0.07	21.98 \pm 0.05	20.65 \pm 0.06	1 \times 1
	1	30.92 \pm 0.06	26.98 \pm 0.07	24.62 \pm 0.07	22.69 \pm 0.08	21.04 \pm 0.08	3 \times 3
	3	35.14 \pm 0.06	30.97 \pm 0.13	28.11 \pm 0.11	25.76 \pm 0.15	23.96 \pm 0.15	
	5	36.04 \pm 0.13	31.91 \pm 0.14	29.05 \pm 0.19	26.55 \pm 0.19	24.77 \pm 0.13	
	7	36.36 \pm 0.16	32.53 \pm 0.18	29.43 \pm 0.22	26.77 \pm 0.15	24.78 \pm 0.15	
	15	36.08 \pm 0.13	32.54 \pm 0.16	29.19 \pm 0.18	26.64 \pm 0.16	24.51 \pm 0.15	
	1	29.97 \pm 0.05	26.17 \pm 0.08	23.96 \pm 0.07	22.18 \pm 0.09	20.82 \pm 0.06	7 \times 7
	3	32.96 \pm 0.08	29.43 \pm 0.08	26.81 \pm 0.10	24.68 \pm 0.12	23.05 \pm 0.10	
	5	34.11 \pm 0.09	30.48 \pm 0.11	27.80 \pm 0.12	25.57 \pm 0.08	23.87 \pm 0.14	
	7	34.62 \pm 0.08	30.81 \pm 0.16	28.30 \pm 0.15	26.18 \pm 0.14	24.42 \pm 0.14	
	15	35.23 \pm 0.09	31.11 \pm 0.10	28.59 \pm 0.13	26.30 \pm 0.18	24.32 \pm 0.20	
	1	31.10 \pm 0.11	24.85 \pm 0.08	22.35 \pm 0.05	20.85 \pm 0.07	19.78 \pm 0.06	3 \times 3
	3	29.07 \pm 0.14	23.64 \pm 0.11	21.04 \pm 0.08	19.88 \pm 0.09	18.73 \pm 0.10	
	5	23.58 \pm 0.06	20.46 \pm 0.02	18.28 \pm 0.04	17.51 \pm 0.06	17.15 \pm 0.06	
	7	20.10 \pm 0.04	18.20 \pm 0.03	17.09 \pm 0.04	15.76 \pm 0.11	15.52 \pm 0.03	
	15	14.36 \pm 0.02	13.74 \pm 0.02	13.29 \pm 0.02	12.81 \pm 0.04	12.76 \pm 0.01	
Non Local PCA compression	1	32.29 \pm 0.08	27.34 \pm 0.08	24.43 \pm 0.11	22.26 \pm 0.16	20.77 \pm 0.14	7 \times 7
	3	33.80 \pm 0.12	28.92 \pm 0.11	25.57 \pm 0.16	22.85 \pm 0.18	20.83 \pm 0.17	
	5	34.05 \pm 0.10	28.98 \pm 0.12	25.46 \pm 0.22	22.47 \pm 0.18	20.33 \pm 0.22	
	7	32.36 \pm 0.12	27.24 \pm 0.14	23.92 \pm 0.13	20.97 \pm 0.15	18.94 \pm 0.14	
	15	25.92 \pm 0.06	20.51 \pm 0.09	17.38 \pm 0.11	15.27 \pm 0.07	8.95 \pm 1.83	
	1	33.81 \pm 0.08	29.10 \pm 0.10	26.34 \pm 0.08	24.48 \pm 0.14	22.89 \pm 0.10	3 \times 3
	3	35.47 \pm 0.11	31.47 \pm 0.16	29.05 \pm 0.15	27.16 \pm 0.15	25.40 \pm 0.17	
	5	35.33 \pm 0.09	31.52 \pm 0.14	29.07 \pm 0.16	27.16 \pm 0.16	25.55 \pm 0.18	
	7	35.51 \pm 0.10	31.73 \pm 0.10	29.04 \pm 0.14	27.27 \pm 0.21	24.25 \pm 0.00	
	15	34.60 \pm 0.08	30.42 \pm 0.09	28.81 \pm 0.15	26.90 \pm 0.18	25.59 \pm 0.13	
	1	32.86 \pm 0.06	28.30 \pm 0.07	25.31 \pm 0.10	23.03 \pm 0.11	21.42 \pm 0.07	7 \times 7
	3	34.79 \pm 0.10	30.16 \pm 0.11	27.36 \pm 0.10	25.18 \pm 0.11	23.48 \pm 0.12	
	5	35.31 \pm 0.07	30.86 \pm 0.11	28.07 \pm 0.14	25.90 \pm 0.14	24.26 \pm 0.17	
	7	35.68 \pm 0.10	31.17 \pm 0.15	28.49 \pm 0.10	26.35 \pm 0.18	24.55 \pm 0.14	
	15	35.99 \pm 0.11	31.64 \pm 0.11	28.53 \pm 0.14	26.37 \pm 0.21	24.78 \pm 0.16	

Note: The largest PSNR values are marked for each method.

numerical instability, which is particularly advantageous to reduce the number of iterations (Correia *et al* 2011).

5. Conclusion

In this study, we proposed a modification to the AD method, motivated by the effectiveness and simplicity of the NLM method for image denoising. The proposed PAD method uses a patch-based approach instead of a local neighbourhood strategy. Additionally, patches are wavelet compressed to speed up the method.

We used a denoising problem to compare our PAD method with the AD method, iterative NLM and its corresponding local model. The results obtained show that combining AD

Table 3. Denoising PSNR results for a test image with various noise levels (Test 1) obtained using AD, PAD, PAD-PCA and PAD-WT.

	\mathcal{W}	10%	20%	30%	40%	50%	P	Time (s)
Local AD	1	29.65 ± 0.04	26.35 ± 0.05	24.06 ± 0.06	22.41 ± 0.09	21.11 ± 0.08	1 × 1	0.11
	1	32.01 ± 0.06	28.21 ± 0.09	25.78 ± 0.08	23.80 ± 0.12	21.94 ± 0.12		0.31
	3	35.66 ± 0.06	31.35 ± 0.12	28.39 ± 0.15	26.03 ± 0.22	24.28 ± 0.18		0.97
	5	36.64 ± 0.10	32.30 ± 0.18	29.27 ± 0.20	26.68 ± 0.20	24.75 ± 0.19	3 × 3	2.19
	7	36.84 ± 0.13	32.61 ± 0.09	29.47 ± 0.19	26.79 ± 0.19	24.83 ± 0.17		3.87
	15	36.84 ± 0.13	32.59 ± 0.14	29.53 ± 0.14	26.34 ± 0.14	24.26 ± 0.21		15.80
	1	30.77 ± 0.04	27.22 ± 0.06	24.95 ± 0.07	22.97 ± 0.09	21.46 ± 0.11		1.13
	3	33.88 ± 0.09	30.02 ± 0.09	27.37 ± 0.08	25.33 ± 0.10	23.65 ± 0.12		2.60
	5	34.42 ± 0.09	31.03 ± 0.11	28.30 ± 0.10	26.22 ± 0.18	24.36 ± 0.11	7 × 7	5.18
	7	34.62 ± 0.07	31.35 ± 0.16	28.65 ± 0.14	26.58 ± 0.20	24.53 ± 0.18		9.21
	15	35.03 ± 0.07	31.51 ± 0.11	28.88 ± 0.12	26.22 ± 0.21	24.79 ± 0.17		37.78
	1	31.30 ± 0.11	25.53 ± 0.11	22.84 ± 0.07	21.27 ± 0.06	20.06 ± 0.07		0.28
	3	32.06 ± 0.10	25.73 ± 0.09	22.87 ± 0.08	20.83 ± 0.09	19.80 ± 0.07		0.85
	5	28.81 ± 0.05	24.08 ± 0.06	20.83 ± 0.05	19.42 ± 0.08	18.38 ± 0.06	3 × 3	1.94
	7	28.35 ± 0.06	22.39 ± 0.06	18.82 ± 0.05	17.79 ± 0.06	16.92 ± 0.06		3.44
PAD PCA compression	15	26.31 ± 0.04	19.05 ± 0.02	16.76 ± 0.05	14.87 ± 0.04	14.24 ± 0.06		13.98
	1	32.41 ± 0.10	27.62 ± 0.10	24.64 ± 0.09	21.68 ± 0.29	17.97 ± 0.58		0.78
	3	33.95 ± 0.11	28.96 ± 0.12	25.54 ± 0.15	22.71 ± 0.17	20.32 ± 0.31		1.72
	5	33.65 ± 0.10	29.08 ± 0.15	25.41 ± 0.21	22.54 ± 0.19	20.37 ± 0.19	7 × 7	3.47
	7	33.57 ± 0.12	28.22 ± 0.17	24.49 ± 0.13	21.63 ± 0.22	19.32 ± 0.16		5.98
	15	26.36 ± 0.10	20.89 ± 0.09	18.43 ± 0.09	16.46 ± 0.10	15.13 ± 0.11		23.51
	1	33.81 ± 0.07	29.07 ± 0.09	26.38 ± 0.08	24.38 ± 0.09	22.63 ± 0.10		0.20
	3	35.88 ± 0.08	31.63 ± 0.12	29.17 ± 0.19	27.10 ± 0.21	25.40 ± 0.22		0.57
	5	36.22 ± 0.15	31.82 ± 0.15	29.22 ± 0.14	27.34 ± 0.16	25.53 ± 0.19	3 × 3	1.21
	7	36.44 ± 0.10	32.05 ± 0.16	29.31 ± 0.16	27.31 ± 0.16	25.62 ± 0.20		2.22
	15	35.47 ± 0.08	31.20 ± 0.14	28.84 ± 0.11	27.15 ± 0.19	25.25 ± 0.18		9.37
	1	32.91 ± 0.06	28.39 ± 0.08	25.49 ± 0.09	23.23 ± 0.10	21.50 ± 0.08		0.63
	3	34.90 ± 0.09	30.34 ± 0.10	27.50 ± 0.13	25.32 ± 0.18	23.71 ± 0.15		1.24
	5	35.54 ± 0.12	31.01 ± 0.12	28.24 ± 0.12	26.11 ± 0.15	24.36 ± 0.17	7 × 7	2.39
	7	35.82 ± 0.08	31.44 ± 0.13	28.62 ± 0.16	26.29 ± 0.17	24.55 ± 0.20		4.01
	15	35.96 ± 0.12	31.72 ± 0.18	28.77 ± 0.13	26.39 ± 0.23	24.35 ± 0.14		15.29

Note: The largest PSNR values are marked for each method.

with NLM, i.e. the PAD method, provides better denoising results than these methods alone. Moreover, results show that the proposed PAD-WT method has superior denoising performance compared to the analogue method with PCA patch compression. In addition, this method is faster and achieves comparable results or, for high noise levels, even better results than those obtained when patch compression is not used.

Medical images are often corrupted by noise and artifacts intrinsic to the image acquisition process. Our study suggests that the proposed method is particularly suitable for medical imaging applications.

In the second part of this paper, we use the PAD-WT as a regularisation method in fDOT image reconstruction. A split operator method is used for solving the fDOT inverse problem,

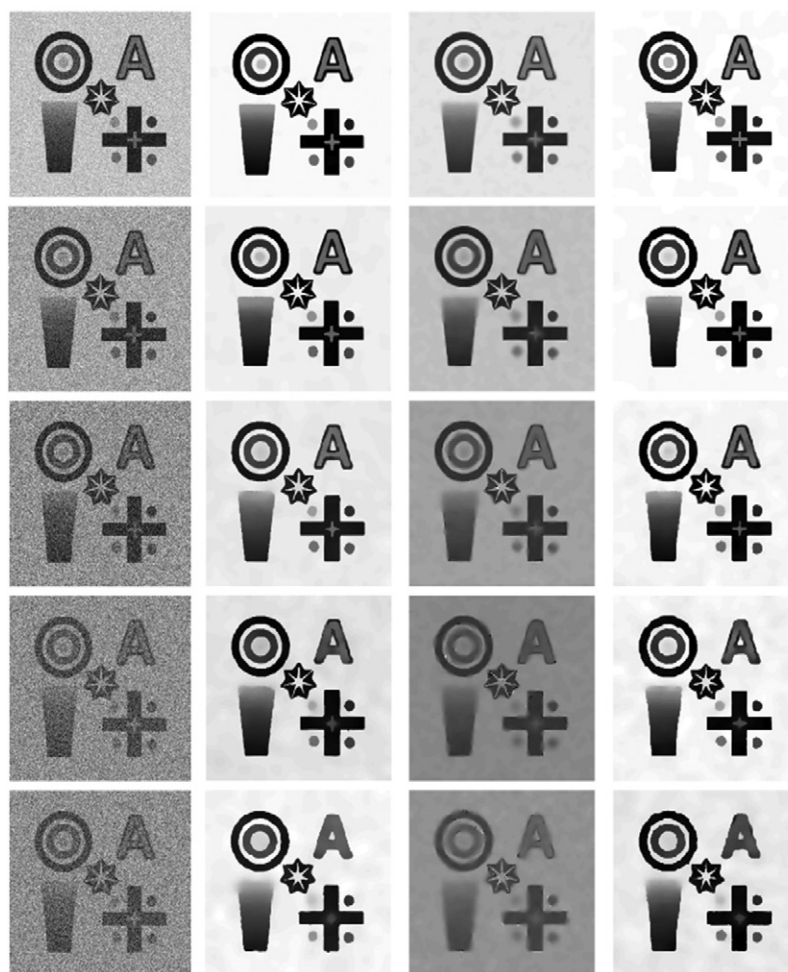


Figure 3. Denoising results of Test 1. (1st column) Test image corrupted with (top to bottom) 10%, 20%, 30%, 40% and 50% Gaussian noise and images denoised using (2nd column) PAD (3rd column) PAD-PCA and (4th column) PAD-WT.

Table 4. Denoising PSNR results for Test 2 images with 20% Gaussian noise obtained using PAD, PAD-PCA and PAD-WT.

	20% noise			
	Barbara	Cameraman	CT	MRI
PAD	26.65 ± 0.03	29.82 ± 0.05	29.17 ± 0.02	31.82 ± 0.05
PAD-PCA	25.99 ± 0.05	29.15 ± 0.06	29.03 ± 0.03	30.92 ± 0.06
PAD-WT	26.43 ± 0.03	29.53 ± 0.03	29.74 ± 0.02	32.11 ± 0.04

which is a two-step method, where a reconstruction step alternates with the regularisation step, i.e. the PAT-WT method.



Figure 4. Denoising results of Test 2. (1st column) Test images (top to bottom) Barbara, cameraman, CT and MRI corrupted with 20% Gaussian noise and images denoised using (2nd column) PAD (3rd column) PAD-PCA and (4th column) PAD-WT.

Acknowledgments

This work has been supported by funding from ECs seventh framework programme FMT-XCT under Grant Agreement No. 201792.

References

- Brox T, Kleinschmidt O and Cremers D 2008 Efficient nonlocal means for denoising of textural patterns *IEEE Trans. Image Process.* **17** 1083–92
- Buades A, Coll B and Morel J-M 2005a A non-local algorithm for image denoising *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition* vol 2 pp 60–5
- Buades A, Coll B and Morel J-M 2005b A review of image denoising algorithms, with a new one *Multiscale Model. Simul.* **4** 490–530
- Chan C, Fulton R, Barnett R, Feng D D and Meikle S 2014 Postreconstruction nonlocal means filtering of whole-body PET with an anatomical prior *IEEE Trans. Med. Imaging* **33** 636–50
- Chen Y, Ma J, Feng Q, Luo L, Shi P and Chen W 2008 Nonlocal prior Bayesian tomographic reconstruction *J. Math. Imaging Vis.* **30** 133–46

- Chen Y, Yang Z, Hu Y, Yang G, Zhu Y, Li Y, Luo L, Chen W and Toumoulin C 2012 Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means *Phys. Med. Biol.* **57** 2667
- Chun S, Fessler J and Dewaraja Y 2012 Non-local means methods using CT side information for I-131 SPECT image reconstruction *IEEE Nuclear Science Symp. Conf. Record* pp 3362–6
- Correia T, Aguirre J, Sisniega A, Chamorro-Servent J, Abascal J, Vaquero J, Desco M, Kolehmainen V and Arridge S 2011 Split operator method for fluorescence diffuse optical tomography using anisotropic diffusion regularisation with prior anatomical information *Biomed. Opt. Express* **2** 2632–48
- Coupé P, Yger P, Prima S, Hellier P, Kervrann C and Barillot C 2008 An optimized blockwise nonlocal means denoising filter for 3D magnetic resonance images *IEEE Trans. Med. Imaging* **27** 425–42
- Coupé P, Manjon J, Robles M and Collins D 2012 Adaptive multiresolution non-local means filter for three-dimensional magnetic resonance image denoising *IET Image Process.* **6** 558–68
- Cuomo S, Michele P and Piccialli F 2014 3D data denoising via nonlocal means filter by using parallel GPU strategies *Comput. Math. Methods Med.* **2014** 523862
- Donoho D 1995 Denoising via soft thresholding *IEEE Trans. Inf. Theory* **41** 613–27
- Donoho D and Johnstone J 1995 Adapting to unknown smoothness via wavelet shrinkage *J. Am. Stat. Assoc.* **90** 1200–24
- Donoho D and Malik J 1994 Ideal spatial adaptation by wavelet shrinkage *Biometrika* **81** 425–55
- Dutta J, Leahy R and Li Q 2013 Non-local means denoising of dynamic PET images *PLoS One* **8** e81390
- Hansen P C 1998 *Rank-Deficient and Discrete ILL-Posed Problems: Numerical Aspects of Linear Inversion* (Philadelphia, PA: SIAM)
- Mahmoudi M and Sapiro G 2005 Fast image and video denoising via nonlocal means of similar neighborhoods *IEEE Signal Process. Lett.* **12** 839–42
- Mallat S 2008 A theory for multiresolution signal decomposition: the wavelet representation *IEEE Trans. Pattern Anal. Mach. Intell.* **11** 674–93
- Nguyen V-G and Lee S-J 2013 Incorporating anatomical side information into PET reconstruction using nonlocal regularization *IEEE Trans. Image Process.* **22** 3961–73
- Ntziachristos V 2006 Fluorescence molecular imaging *Ann. Rev. Biomed. Eng.* **8** 1–33
- Orchard J, Ebrahimi M and Wong A 2008 Efficient nonlocal-means denoising using the SVD *IEEE Int. Conf. Image Process.* pp 1732–5
- Perona P and Malik J 1990 Scale-space and edge detection using anisotropic diffusion *IEEE Trans. Pattern Anal. Mach. Intell.* **12** 629–39
- Salomon D and Motta G 2009 *Handbook of Data Compression* 5th edn (Berlin: Springer)
- Tasdizen T 2008 Principal components for non-local means image denoising *IEEE Int. Conf. Image Process.* pp 1728–31
- Wang G and Qi J 2012 Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization *IEEE Trans. Med. Imaging* **31** 2194–204
- Zacharopoulos A, Svenmarker P, Axelsson J, Schweiger M, Arridge S and Andersson-Engels S 2009 A matrix-free algorithm for multiple wavelength fluorescence tomography *Opt. Express* **17** 3025–35